



Complex Adaptive Systems and the Evolution of Reciprocation

Sigmund, K.

IIASA Interim Report
December 1998



Sigmund, K. (1998) Complex Adaptive Systems and the Evolution of Reciprocation. IIASA Interim Report. Copyright © 1998 by the author(s). <http://pure.iiasa.ac.at/5552/>

Interim Report on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work. All rights reserved. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage. All copies must bear this notice and the full citation on the first page. For other purposes, to republish, to post on servers or to redistribute to lists, permission must be sought by contacting repository@iiasa.ac.at

INTERIM REPORT IR-98-100 / December

Complex Adaptive Systems and the Evolution of Reciprocation

Karl Sigmund (ksigmund@esi.ac.at)

Approved by
Ulf Dieckmann (dieckman@iiasa.ac.at)
Project Coordinator, *Adaptive Dynamics Network*



The Adaptive Dynamics Network at IIASA fosters the development of new mathematical and conceptual techniques for understanding the evolution of complex adaptive systems.

Focusing on these long-term implications of adaptive processes in systems of limited growth, the Adaptive Dynamics Network brings together scientists and institutions from around the world with IIASA acting as the central node.

Scientific progress within the network is reported in the IIASA Studies in Adaptive Dynamics series.

THE ADAPTIVE DYNAMICS NETWORK

The pivotal role of evolutionary theory in life sciences derives from its capability to provide causal explanations for phenomena that are highly improbable in the physicochemical sense. Yet, until recently, many facts in biology could not be accounted for in the light of evolution. Just as physicists for a long time ignored the presence of chaos, these phenomena were basically not perceived by biologists.

Two examples illustrate this assertion. Although Darwin's publication of "The Origin of Species" sparked off the whole evolutionary revolution, oddly enough, the population genetic framework underlying the modern synthesis holds no clues to speciation events. A second illustration is the more recently appreciated issue of jump increases in biological complexity that result from the aggregation of individuals into mutualistic wholes.

These and many more problems possess a common source: the interactions of individuals are bound to change the environments these individuals live in. By closing the feedback loop in the evolutionary explanation, a new mathematical theory of the evolution of complex adaptive systems arises. It is this general theoretical option that lies at the core of the emerging field of adaptive dynamics. In consequence a major promise of adaptive dynamics studies is to elucidate the long-term effects of the interactions between ecological and evolutionary processes.

A commitment to interfacing the theory with empirical applications is necessary both for validation and for management problems. For example, empirical evidence indicates that to control pests and diseases or to achieve sustainable harvesting of renewable resources evolutionary deliberation is already crucial on the time scale of two decades.

The Adaptive Dynamics Network has as its primary objective the development of mathematical tools for the analysis of adaptive systems inside and outside the biological realm.

IIASA STUDIES IN ADAPTIVE DYNAMICS

- No. 1 Metz JAJ, Geritz SAH, Meszéna G, Jacobs FJA, van Heerwaarden JS:
Adaptive Dynamics: A Geometrical Study of the Consequences of Nearly Faithful Reproduction.
IIASA Working Paper WP-95-099.
van Strien SJ, Verduyn Lunel SM (eds.): Stochastic and Spatial Structures of Dynamical Systems, Proceedings of the Royal Dutch Academy of Science (KNAW Verhandeligen), North Holland, Amsterdam, pp. 183-231 (1996).
- No. 2 Dieckmann U, Law R:
The Dynamical Theory of Coevolution: A Derivation from Stochastic Ecological Processes.
IIASA Working Paper WP-96-001.
Journal of Mathematical Biology (1996) 34, 579–612.
- No. 3 Dieckmann U, Marrow P, Law R:
Evolutionary Cycling of Predator-Prey Interactions: Population Dynamics and the Red Queen.
Journal of Theoretical Biology (1995) 176, 91–102.
- No. 4 Marrow P, Dieckmann U, Law R:
Evolutionary Dynamics of Predator-Prey Systems: An Ecological Perspective.
IIASA Working Paper WP-96-002.
Journal of Mathematical Biology (1996) 34, 556–578.
- No. 5 Law R, Marrow P, Dieckmann U:
On Evolution under Asymmetric Competition.
IIASA Working Paper WP-96-003.
Evolutionary Ecology (1997) 11, 485–501.
- No. 6 Metz JAJ, Mylius SD, Dieckmann O:
When Does Evolution Optimise? On the Relation between Types of Density Dependence and Evolutionarily Stable Life History Parameters.
IIASA Working Paper WP-96-004.
- No. 7 Ferrière R, Gatto M:
Lyapunov Exponents and the Mathematics of Invasion in Oscillatory or Chaotic Populations.
Theoretical Population Biology (1995) 48, 126–171.
- No. 8 Ferrière R, Fox GA:
Chaos and Evolution.
Trends in Ecology and Evolution (1995) 10, 480–485.
- No. 9 Ferrière R, Michod RE:
The Evolution of Cooperation in Spatially Heterogeneous Populations.
IIASA Working Paper WP-96-029.
American Naturalist (1996) 147, 692–717.

- No. 10 Van Dooren TJM, Metz JAJ:
Delayed Maturation in Temporally Structured Populations with Non-Equilibrium Dynamics.
IIASA Working Paper WP-96-070.
Journal of Evolutionary Biology (1998) 11, 41–62.
- No. 11 Geritz SAH, Metz JAJ, Kisdi E, Meszéna G:
The Dynamics of Adaptation and Evolutionary Branching.
IIASA Working Paper WP-96-077.
Physical Review Letters (1997) 78, 2024–2027.
- No. 12 Geritz SAH, Kisdi E, Meszéna G, Metz JAJ:
Evolutionarily Singular Strategies and the Adaptive Growth and Branching of the Evolutionary Tree.
IIASA Working Paper WP-96-114.
Evolutionary Ecology (1998) 12, 35–57.
- No. 13 Heino M, Metz JAJ, Kaitala V:
Evolution of Mixed Maturation Strategies in Semelparous Life-Histories: the Crucial Role of Dimensionality of Feedback Environment.
IIASA Working Paper WP-96-126.
Philosophical Transactions of the Royal Society of London Series B (1997) 352, 1647–1655.
- No. 14 Dieckmann U:
Can Adaptive Dynamics Invade?
IIASA Working Paper WP-96-152.
Trends in Ecology and Evolution (1997) 12, 128–131.
- No. 15 Meszéna G, Czibula I, Geritz SAH:
Adaptive Dynamics in a Two-Patch Environment: a Simple Model for Allopatric and Parapatric Speciation.
IIASA Interim Report IR-97-001.
Journal of Biological Systems (1997) 5, 265–284.
- No. 16 Heino M, Metz JAJ, Kaitala V:
The Enigma of Frequency-Dependent Selection.
IIASA Interim Report IR-97-061.
Trends in Ecology and Evolution (1998) in press.
- No. 17 Heino M:
Management of Evolving Fish Stocks.
IIASA Interim Report IR-97-062.
Canadian Journal of Fisheries and Aquatic Sciences (1998) in press.
- No. 18 Heino M:
Evolution of Mixed Reproductive Strategies in Simple Life-History Models.
IIASA Interim Report IR-97-063.
- No. 19 Geritz SAH, van der Meijden E, Metz JAJ:
Evolutionary Dynamics of Seed Size and Seedling Competitive Ability.
IIASA Interim Report IR-97-071.

- No. 20 Galis F, Metz JAJ:
Why are there so many Cichlid Species? On the Interplay of Speciation and Adaptive Radiation.
IIASA Interim Report IR-97-072.
Trends in Ecology and Evolution (1998) 13, 1–2.
- No. 21 Boerlijst MC, Nowak MA, Sigmund K:
Equal Pay for all Prisoners. / The Logic of Contrition.
IIASA Interim Report IR-97-073.
AMS Monthly (1997) 104, 303–307.
Journal of Theoretical Biology (1997) 185, 281–294.
- No. 22 Law R, Dieckmann U:
Symbiosis without Mutualism and the Merger of Lineages in Evolution.
IIASA Interim Report IR-97-074.
Proceedings of the Royal Society of London Series B (1998) 265, 1245–1253.
- No. 23 Klinkhamer PGL, de Jong TJ, Metz JAJ:
Sex and Size in Cosexual Plants.
IIASA Interim Report IR-97-078.
Trends in Ecology and Evolution (1997) 12, 260–265.
- No. 24 Fontana W, Schuster P:
Shaping Space: The Possible and the Attainable in RNA Genotype-Phenotype Mapping.
IIASA Interim Report IR-98-004.
- No. 25 Kisdi E, Geritz SAH:
Adaptive Dynamics in Allele Space: Evolution of Genetic Polymorphism by Small Mutations in a Heterogeneous Environment.
IIASA Interim Report IR-98-038.
- No. 26 Fontana W, Schuster P:
Continuity in Evolution: On the Nature of Transitions.
IIASA Interim Report IR-98-039.
Science (1998) 280, 1451–1455.
- No. 27 Nowak MA, Sigmund K:
Evolution of Indirect Reciprocity by Image Scoring. / The Dynamics of Indirect Reciprocity.
IIASA Interim Report IR-98-040.
Nature (1998) 393, 573–577.
- No. 28 Kisdi E:
Evolutionary Branching Under Asymmetric Competition.
IIASA Interim Report IR-98-045.
- No. 29 Berger U:
Best Response Adaptation for Role Games.
IIASA Interim Report IR-98-086.
- No. 30 Van Dooren TJM:
The Evolutionary Ecology of Dominance-Recessivity
IIASA Interim Report IR-98-096.

- No. 31 Dieckmann U, O'Hara B, Weisser W:
The Evolutionary Ecology of Dispersal.
IIASA Interim Report IR-98-108.
Trends in Ecology and Evolution (1999) in press.
- No. 32 Sigmund K:
Complex Adaptive Systems and the Evolution of Reciprocation.
IIASA Interim Report IR-98-100.
- No. 33 Posch M, Pichler A, Sigmund K:
The Efficiency of Adapting Aspiration Levels.
IIASA Interim Report IR-98-103.

Issues of the IIASA Studies in Adaptive Dynamics series can be obtained free of charge. Please contact:

Adaptive Dynamics Network
International Institute for Applied Systems Analysis
Schloßplatz 1
A-2361 Laxenburg
Austria

Telephone +43 2236 807, Telefax +43 2236 71313, E-Mail adn@iiasa.ac.at,
Internet <http://www.iiasa.ac.at/Research/ADN>

Contents

1	Introduction	1
2	Direct Reciprocation	2
3	Indirect Reciprocation	5
4	Discussion	6

About the Author

Karl Sigmund
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Vienna, Austria
and
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

Complex Adaptive Systems and the Evolution of Reciprocation

Karl Sigmund

1 Introduction

Most of the major transitions in evolution consist in assembling units of some type to form a new, hierarchically higher entity (Maynard Smith and Szathmary, 1995). As a rule, the reproductive potential is monopolised by a minority of the previous units. This is most notably the case with multi-cellular organisms, where only germ-line cells transmit the genome, or with colonies of eusocial animals having large proportions of sterile workers. This channeling of reproductive opportunities entails a close relatedness between elements of the higher order unit, and allows for cooperation based on kin selection.

But in human societies, where cooperation is ubiquitous, there are remarkably few individual inequalities in reproductive potential. A few potentates have managed to obtain an almost unlimited control of their community and to sire several hundred offspring, but these are exceptions which occurred at a late, and probably transient, stage of cultural evolution. Both in modern mass societies and in bands of hunter-gatherers, social rules tend to level reproductive opportunities and to prevent the establishment of a global controller. While a large part of the services and tasks performed within households can be explained by kin selection, the bulk of human cooperation is based, not on relatedness, but on reciprocation. Not genetic ties, but economic exchanges explain the cohesion of human communities.

This implies that individuals collaborate only if it is to their own good. The mathematical framework for studying the economics of interacting egoists is game theory: more precisely, that branch of game theory which is, somewhat misleadingly, called non-cooperative game theory (cf. Binmore, 1994). The term ‘non-cooperative’ means in this context that players cannot negotiate binding and enforceable agreements. Such players can nevertheless achieve cooperation, not by the command of a controller or by deference to the benefit of the group, but by following a myopic set of rules evolved to optimise their selfish interests. The challenge lies in showing how they do it in spite of the ever-present lure of defection.

Robert Trivers was the first to suggest *reciprocation* as a basis for mutual assistance in animal behaviour (including human behaviour), and to discuss it in terms of game theory. More precisely, he introduced the Prisoner’s Dilemma game to bring the problem into focus (Trivers, 1971; see also May, 1987, and Boyd, 1988). This approach was greatly expanded by Robert Axelrod and William D. Hamilton (1981), who applied evolutionary game theory (see Maynard Smith, 1982) to populations of

players engaged in many rounds of the repeated Prisoner’s Dilemma game against randomly chosen co-players. Already at this early stage, computer simulations were essential to follow the complex nonlinear dynamics of the frequencies of strategies submitted to natural selection (Axelrod, 1984).

A further decisive step was taken when Axelrod adapted the genetic algorithms of John Holland (1975) to simulate the effects of evolutionary trial and error. This was the first application of genetic algorithms to a genuine evolutionary problem (rather than a technical optimization problem). It proved particularly successful because the binary coding of strategies, usually a moot point with genetic algorithms, was straightforward (Axelrod, 1987, see also Axelrod, 1997, and Epstein and Axtell, 1996).

This success led to many further investigations, including the effects of noise, extending the memory, allowing for alternating moves, introducing more complex state-based strategies, increasing the number of players, changing the payoff structure, restricting the interactions to neighbours only, addressing continuous versions of the game etc. We present a short survey of this field, with pointers to the biological motivations behind these variants. The basic message is that details matter, but that cooperation robustly emerges from a bottom-up approach based on reciprocity. We finally deal with indirect reciprocity, which is, according to Richard Alexander (1987), the biological basis of human moral systems. In this setup, an act of assistance may be returned, not to the donor, but to a third party. A model based on the status of the players shows that cooperation can emerge even if any two individuals never interact more than once.

This brief survey concentrates on papers dealing with pairwise interactions and using evolutionary simulations. This is certainly not intended to deny the importance of other approaches to the evolution of cooperation (see Dugatkin et al, 1992).

2 Direct Reciprocation

The Prisoner’s Dilemma (PD) is a two-player game where both players have the same two strategies and the same payoffs. The two strategies are **C** (to cooperate) and **D** (to defect). If both players use **C**, both get the *reward* R for mutual cooperation; if both play **D**, both get the *penalty* P for mutual defection; a **D** player obtains the *temptation* payoff T for unilaterally defecting against a co-player who plays **C**, whereas the co-player receives the *sucker’s* payoff S for being exploited. One assumes $T > R > P > S$ and $2R > T + S$. The first condition implies that **D** dominates **C**, in the sense that it is better no matter what the other player chooses, and the second condition entails that joint cooperation is better than sharing the payoffs after a unilateral defection. Players have to chose **D** if they want to maximise their payoff, and this yields the payoff P , which is less than the reward R for mutual cooperation.

In the repeated PD game, players have to chose simultaneously, in every round, whether to play **C** or **D**. There is a constant probability w for another round, so that the average length of the game is random. The total payoff is given by $\sum A_n w^n$, where A_n is the payoff in round n . The strategies for the repeated PD can be arbitrarily complex, but in a biological context, it makes only sense to consider

strategies given by simple knee-jerk rules, like Tit For Tat (TFT, which plays **C** in the first round and then simply repeats the co-players previous move). The first lesson from Axelrod’s computer tournaments was that such simple rules, and in particular TFT, more than hold their own against more sophisticated strategies (based on stochastic optimisation, for instance).

Complex adaptive systems are used to simulate artificial societies of players engaged in repeated PD games. In Axelrod (1987), strategies were considered which played **C** or **D** depending on the outcome of the previous three rounds. They were coded by binary strings of length 70. For every new generation, these strings were submitted to point-mutation and recombination. The players then engaged in a round-robin tournament, accumulating payoff points which translated into number of offspring to form the next generation. An initial increase in defectors was often followed by a re-emergence of cooperation.

Bendor et al (1991), Nowak and Sigmund (1992) and Kollock (1993) have considered the influence of noise, which is particularly disruptive in a society dominated by TFT. When one allows only for reactive strategies given by different propensities to play **C**, depending on the last move of the co-player, one finds that cooperation emerges based on a tolerant rule called GTFT (Generous TFT), which retaliates only with a certain probability after a **D**, but always cooperates after a **C** (up to mistakes in implementation). This rule cannot spread in a society of defectors, however; it is necessary first that TFT invades and paves the way, like a pioneering species in a plant community, for GTFT to take over (Nowak and Sigmund, 1992).

If one considers strategies which depend on the moves of both players in the previous round, one finds a rich collection of far-from-equilibrium dynamics even if only three or four strategies interact (Nowak and Sigmund, 1993a). Sufficiently long mutation-selection chronicles usually lead to meta-stable cooperative regimes, however, based (in case $P + T < 2R$) on Pavlov, a rule which cooperates if both players used the same move in the previous round. Pavlov is a win-stay, lose-shift rule: players repeat their previous move if it led to a high payoff (R or T), and try the other option if the payoff was low (S or P). In a population of Pavlov players, unilateral defections due to mistakes cause one round of mutual defection, after which both players resume cooperation (Nowak and Sigmund, 1993b, Milinski, 1993). Pavlov exploits unconditional cooperators, preventing them from invading and thus offering targets for exploiters. Again, cooperation can only emerge after the invasion of stern retaliatory strategies like TFT, which then are superseded by the error-proof Pavlov. TFT acts in this sense as a catalyser for cooperation.

It should be stressed that these simulations of artificial populations based on extended mutation-selection chronicles show a high degree of history-dependence and often display punctuated equilibrium. This is particularly clear in Lindgren (1991), where extensions of the memory were possible (caused by mutations reminiscent of gene duplications which introduce strategies based not just the last round, but on a larger window of the past). Usually such simulations led to the emergence of cooperative strategies similar to Pavlov, or of variants defecting twice in a row, after a mistake.

There exist very simple finite-state automata which cannot be described by rules depending only on the outcome of a given number of previous rounds. An example is

Contrite TFT (CTFT), a strategy that monitors the own standing and that of the co-player. A player's standing is good except after defecting against a player with good standing. CTFT-players cooperate except if they are in good standing and their co-player is in bad standing. In Boerlijst et al (1997) it is shown that such strategies are good at invading populations of defectors and establishing a stable cooperative regime. However, they are only immune against errors in implementation, whereas Pavlov is also proof against errors in perception. Leimar (1997) has shown that there exists a huge number of finite-state automata leading to limit-ESS (a version of evolutionarily stable strategies). This implies a pronounced path-dependence of evolutionary chronicles.

In the usual PD setup, both players are supposed to move at once. This is the case, for instance, in predator-inspection games by sticklebacks or guppies, where the fish take their mirror-image for a co-player (Milinski, 1987). In many situations of mutual aid, however, players move alternately. This is the case, for instance, when a well-fed vampire bat feeds a hungry conspecific (Wilkinson, 1984), or when a young male baboon diverts the attention of the dominant male so that his pal can mount an oestrous female (Packer, 1977). Simulations of the alternating PD (Nowak and Sigmund, 1994, Frean, 1994, Hauert and Schuster, 1998) lead again to cooperation based on error-proof strategies (not Pavlov, however, but Firm But Fair, a strategy that defects only after an unwarranted defection by the co-player). Experiments by Wedekind and Milinski (1997) support this distinction between strategies for the simultaneous and the alternating PD.

If players do not interact at random, but only within a certain neighborhood structure, cooperation becomes much more readily established, even for the one-shot PD (see Nowak and May, 1992 and 1993, Sigmund, 1992, Huberman and Glance, 1993, Nowak et al, 1994a,b, Lindgren and Nordahl, 1994, Killingback and Doebeli, 1996, and for a general introduction to cellular automata in this context, Durrett and Levin, 1994). Again, if players have the possibility of choice and refusal of partners, cooperation becomes much easier to achieve (Stanley, Ashlock and Tesfatsion, 1994). In both cases, aggregation of cooperators occurs through local rules. In contrast, cooperation becomes much more unlikely if more than two players engage in the interaction. For extensive simulations of this setup we refer to Hauert and Schuster (1997).

It should be stressed that the PD is certainly not the only game modelling aspects of cooperation. In particular, Sugden (1987) has stressed that the payoff rank ordering $T > R > S > P$ makes also sense in this context. In such an interaction a player faced with a defector would nevertheless prefer to play **C**. This yields what game theorists call the Chicken game and biologists Hawk-Dove. Repeated Chicken need not lead to mutual cooperation based on reciprocity; it is more likely that an asymmetry becomes effective and decides that one player will always play **C** and the other **D**. This could shed some light on the intruder-inspections by female lions, where one can often distinguish leaders and laggards (see Heinsohn and Packer, 1996).

3 Indirect Reciprocation

In addition to reciprocation based on repeated interactions within a pair, there exists another, indirect reciprocity, where the donor does not obtain a return from the recipient, but from a third party. Donors provide help if the recipient has helped others in the past. This works if the cost of an altruistic act is offset by a raised ‘score’, or status, which increases the chance to subsequently become the recipient of an altruistic act. Cooperation is channelled towards the ‘valuable’ members of the community. For Richard Alexander, ‘indirect reciprocity involves reputation and status, and results in everyone in the group continually being assessed and re-assessed’.

In Nowak and Sigmund (1998) this is modelled by a population of individuals having the options of helping another or not. In each generation, a number of potential donor-recipient pairs are chosen randomly: if the help is actually provided, this implies a cost c to the donor, a benefit b to the recipient, and it increases the donor’s score by one. The score of a player refusing to help is decreased by one. Initially all scores are zero. We consider strategies given by integers k ; a player with such a strategy helps if and only if the score of the potential recipient is at least k . We can follow the frequencies of the strategies from generation to generation, allowing for occasional mutations.

A remarkably small number of interactions (for $b = 10$ and $c = 1$, an average of two interactions per lifetime suffices, for instance) can lead to the emergence of cooperative populations where most members use $k = 0$ or $k = -1$. If the simulation is continued, strategies which are less discriminating spread: players with $k = -3$, for instance, will rarely ever refuse to help, their score will therefore increase faster than average, and hence they will in turn be helped more often. But if the frequency of less discriminating players reaches a certain threshold, then defectors (players with $k = 3$, for instance, who hardly ever provide help) can take over, with the result that cooperation disappears in the population. Once this happens, the average k -values will decrease again, leading eventually back to a cooperative regime of players with maximal discrimination (i.e. $k = 0$). To summarize, random drift can subvert populations of discriminate altruists by indiscriminate altruists; once their frequency is large, defectors can invade; but as soon as the defectors have reduced the proportion of indiscriminate altruists, the discriminate altruists can fight back and eliminate the defectors. This leads again to a cooperative population which is proof against defectors, but not against indiscriminate altruists, etc.

Such models show that indirect reciprocity based on image scoring works in principle. It should be stressed that the chance of two players ever meeting again is vanishingly small. All that needs to be known is the score of the co-player (see also Pollock and Dugatkin, 1992). Even this image scoring need not be public knowledge. One can modify the model to include, for every player, a private assessment of the other group members. In particular, one can assume that an interaction between two individuals is only observed by a small subset of the population. Only these on-lookers will update their score of the donor. For larger groups, it becomes more difficult to establish cooperation.

Models which are even more simplified help to explain analytically the cycling behaviour, with its long bouts of cooperation interspersed by short periods of de-

fection, which is reminiscent of the lack of stability near a critical state. Somewhat surprisingly, cooperation is more robust if the society is challenged more frequently by invasion attempts of defectors. One can compute the minimal amount of discriminators, the minimal number of rounds per generation and the maximal size of the society, for indirect reciprocity to work. This yields as necessary condition for cooperation that the *degree of acquaintanceship* (the probability that a player knows the score of the co-player) is larger than the cost-to-benefit ratio c/b . This result is analogous to Hamilton’s rule which states that the *degree of relatedness* (the probability that an allele in the player’s genome is also present in the co-player) must exceed c/b .

4 Discussion

The success in analysing the iterated PD should not hide the fact that variants with more than 2 players lead only rarely to a cooperative outcome (see Hauert and Schuster, 1997). Such N-person games have attracted much attention, for instance as the Free Rider Problem or as the Tragedy of the Commons (Hardin, 1968). The latter name suggests already, as a major application, the management of ecosystems.

More generally, how can individual restraint in the exploitation of a common resource emerge and establish itself in the absence of a global controller? This problem underlies the evolution of virulence, or the concept of the prudent predator. In the human context, or more generally among individuals recognising each other, the obvious solution – the punishment of selfish exploiters – leads to another dilemma. To punish defectors is a costly, possibly dangerous activity, and it is tempting to leave it to others. But this constitutes a second-order defection, which ought also to be punished, etc. Several authors (Glance and Huberman, Boyd and Richerson 1988) have addressed this problem. It seems plausible that the idea of indirect reciprocity can be applied in this context.

References

- Alexander, R.D. (1987) *The Biology of Moral Systems*, Aldine de Gruyter, New York.
- Axelrod, R. (1984) *The Evolution of Cooperation*, reprinted 1989 by Penguin, Harmondsworth.
- Axelrod, R. (1987) The evolution of strategies in the iterated Prisoner’s Dilemma. In *Genetic algorithms and simulated annealing* (ed. D. Lawrence), p.32-41, Pitman, London.
- Axelrod, R. (1997) *The Complexity of Cooperation*, Princeton UP, Princeton.
- Axelrod, R. and Hamilton, W.D. (1981) The evolution of cooperation, *Science* **211**, 1390-6.
- Bendor, J., Kramer, R.M. and Stout, S. (1991) When in doubt... Cooperation in a noisy Prisoner’s Dilemma, *J. of Conflict Resolution* **35**, 691-719.

- Binmore, K.G. (1994) *Fun and Games*, Heath and Co, Lexington, Mass.
- Boerlijst, M., Nowak, M.A. and Sigmund, K. (1997) The Logic of Contrition, *Journ. Theor. Biol.*, **185**, 281-293.
- Boyd, R. (1988) Is the repeated Prisoner's Dilemma a good model of reciprocal altruism?, *Ethology and Sociobiology* **9**, 211-222.
- Boyd, R. and Richerson, P. (1988) The revolution of reciprocity in sizable groups, *J. Theor. Biol.* **132**, 337-56.
- Dugatkin, L.A., Mesterton-Gibbons, M. and Houston, A.I. (1992) Beyond the Prisoner's Dilemma: toward models to discriminate among mechanisms of cooperation in nature, *TREE* **7**, 202-5.
- Durrett, R. and Levin, S.A. (1994) The importance of being discrete (and spatial), *Theor. Popul. Biol.* **46**, 363-394.
- Epstein, J.M. and Axtell, R. (1996) *Growing Artificial Societies*, The Brookings Institution Press, Mass.
- Frean, M.R. (1994) The Prisoner's Dilemma without synchrony, *Proc. Roy. Soc. London B*, **257**, 75-79.
- Glance, N.S. and Huberman, B.A. (1994) The dynamics of social dilemmas, *Scientific American*, **270** (March), 76-81.
- Hardin, G. (1968) The tragedy of the commons, *Science* , **162**, 1243-8.
- Hauert, Ch. and Schuster, H.G. (1997) Effects of increasing the number of players and memory size in the iterated Prisoner's Dilemma; a numerical approach, *Proc. R. Soc. London B*, **264**, 513-519.
- Hauert, Ch. and Schuster, H.G. (1998) Extending the iterated Prisoner's Dilemma without synchrony, to appear.
- Heinsohn, R. and Packer, C. (1995) Complex cooperative strategies in group-territorial African lions, *Science*, **269**, 1260-2.
- Holland, J. (1975) *Adaptation in Natural and Artificial Systems*, Ann Arbor, Univ. of Michigan Press.
- Hubermann, B.A. and Glance, N.S. (1993) Evolutionary games and computer simulations, *Proc. Nat. Acad. Sci. USA* **90**, 7712-15.
- Killingback, T. and Doebeli, M. (1996) Spatial evolutionary game theory: Hawks and Doves revisited, *Proc. Royal Society London B* **263**, 1135-1144.
- Kollock, P. (1993) An eye for an eye leaves everyone blind, *American Sociological Review* **58**, 768-86.
- Leimar, O. (1997) Repeated games: a state space approach, *Journ. Theor. Biol.*, **184**, 471-98.
- Lindgren, K. (1991) Evolutionary phenomena in simple dynamics, in: *Artificial Life II*, ed. C.G. Langton et al, Addison-Wesley, Redwood City, pp. 295-312.
- Lindgren, K. and Nordahl, M.G. (1994) Evolutionary dynamics of spatial games, *Physica D* **75**, 292-309.

- May, R.M. (1987) More evolution of cooperation, *Nature* **327**, 15-17.
- Maynard Smith, J. (1982) *Evolution and the theory of games*, Cambridge UP.
- Maynard Smith, J. and Szathmary, E. (1995) *The major transitions in evolution*, Freeman, Oxford.
- Milinski, M. (1987) Tit For Tat in stickleback and the evolution of cooperation, *Nature*, **325**, 434-35.
- Milinski, M. (1993) Cooperation wins and stays, *Nature* **364**, 12-13.
- Nowak, M.A. and May, R.M. (1992) Evolutionary Games and Spatial Chaos, *Nature* **359**, 826-29.
- Nowak, M.A. and May, R.M. (1993) The spatial dilemmas of evolution, *Int. J. Bifurcation and Chaos* **3**, 35-78.
- Nowak, M.A. and Sigmund, K. (1992) Tit for Tat in heterogeneous populations, *Nature* **355**, 250-253.
- Nowak, M.A. and Sigmund, K. (1993a) Chaos and the evolution of cooperation, *Proc. Nat. Acad. Sci. USA*, **90**, 5091-94.
- Nowak, M.A. and Sigmund, K. (1993b) Win-stay, lose-shift outperforms tit-for-tat, *Nature*, **364**, 56-8.
- Nowak, M.A. and K. Sigmund (1994) The Alternating Prisoner's Dilemma, *J. theor. Biol.* **168**, 219-226.
- Nowak, M.A., Bonhoeffer, S. and May, R.M. (1994a) Spatial games and the maintenance of cooperation, *Proc. Natl. Acad. Sci.* **91**, 4877-81.
- Nowak, M.A., Bonhoeffer, S. and May, R.M. (1994b) More spatial games, *Int. J. Bifurcation and Chaos*, **4**, 33-56.
- Nowak, M.A. and Sigmund, K. (1998) Evolution of indirect reciprocity by image scoring, to appear in *Nature*.
- Nowak, M.A. and Sigmund, K. (1998) The dynamics of indirect reciprocity, submitted.
- Packer, C. (1977) Reciprocal altruism in *Papio anubis*, *Nature* **265**, 441-3.
- Pollock, G.B. and Dugatkin, L.A. (1992) Reciprocity and the evolution of reputation, *Journ. Theor. Biol.* **159**, 25-37.
- Sigmund, K. (1992) On Prisoners and Cells, *Nature* **359**, 774.
- Stanley, E.A., Ashlock, D. and Tesfatsion, L. (1994) Iterated Prisoner's Dilemma with choice and refusal of partners, in: *Artificial Life III*, ed. C.G. Langton, Addison-Wesley, Redwood City.
- Sugden, R. (1986) *The Evolution of Rights, Co-operation and Welfare*. Blackwell, Oxford.
- Trivers, R. (1971) The evolution of reciprocal altruism, *Quart. Rev. Biol.* **46**, 35-57.
- Wedekind, C. and Milinski, M. (1996) Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus generous tit-for-tat. *Proc. Nat. Acad. Sci. USA*, **93**, 2686-89.
- Wilkinson, G.S. (1984) Reciprocal food-sharing in the vampire bat, *Nature* **308**, 181-4.